PHom-GeM: Persistent Homology for Generative Models

Jeremy Charlier
University of Luxembourg
Luxembourg, Luxembourg
jeremy.charlier@uni.lu

Radu State
University of Luxembourg
Luxembourg, Luxembourg
radu.state@uni.lu

Jean Hilger
BCEE
Luxembourg, Luxembourg
j.hilger@bcee.lu

***Abstract*—Generative neural network models, including Generative Adversarial Network (GAN) and Auto-Encoders (AE), are among the most popular neural network models to generate adversarial data. The GAN model is composed of a generator that produces synthetic data and of a discriminator that discriminates between the generator's output and the true data. AE consist of an encoder which maps the model distribution to a latent manifold and of a decoder which maps the latent manifold to a reconstructed distribution. However, generative models are known to provoke chaotically scattered reconstructed distribution during their training, and consequently, incomplete generated adversarial distributions. Current distance measures fail to address this problem because they are not able to acknowledge the shape of the data manifold, i.e. its topological features, and the scale at which the manifold should be analyzed. We propose Persistent Homology for Generative Models, PHom-GeM, a new methodology to assess and measure the distribution of a generative model. PHom-GeM minimizes an objective function between the true and the reconstructed distributions and uses persistent homology, the study of the topological features of a space at different spatial resolutions, to compare the nature of the true and the generated distributions. Our experiments underline the potential of persistent homology for Wasserstein GAN in comparison to Wasserstein AE and Variational AE. The experiments are conducted on a real-world data set particularly challenging for traditional distance measures and generative neural network models. PHom-GeM is the first methodology to propose a topological distance measure, the bottleneck distance, for generative models used to compare adversarial samples in the context of credit card transactions.**

***Index Terms*—Neural Networks, Optimal Transport, Algebraic Topology**

I. Motivation

The field of unsupervised learning has evolved significantly for the past few years thanks to adversarial networks publications. In [1], Goodfelow et al. introduced a Generative Adversarial Network framework called GAN. It is a class of generative models that play a competitive game between two networks in which the generator network must compete against an adversary according to a game theoretic scenario [2]. The generator network produces samples from a noise distribution and its adversary, the discriminator network, tries to distinguish real samples from generated samples, respectively samples inherited from the training data and samples produced by the generator. Meanwhile, Variational Auto-Encoders (VAE) presented by Kingma et al. in [3] have emerged as a well-established approach for synthetic data generation. Nevertheless, they might generate poor target distribution because of the KL divergence [2]. We recall an AE is a neural network trained to copy its input manifold to its output manifold through a hidden layer. The encoder function sends the input space to the hidden space and the decoder function brings back the hidden space to the input space. By applying some of the Optimal Transport (OT) concepts gathered in [4] and noticeably, the Wasserstein distance, Arjovsky et al. introduced the Wasserstein GAN (WGAN) in [5]. It tries to avoid the mode collapse, a typical training convergence issue occurring between the generator and the discriminator. Gulrajani et al. further optimized the concept in [6] by proposing a Gradient Penalty to Wasserstein GAN (GP-WGAN) capable to generate adversarial samples of higher quality. Similarly, Tolstikhin et al. in [7] applied the same OT concepts to AE and, therefore, introduced Wasserstein AE (WAE), a new type of AE generative model, that avoids the use of the KL divergence.

Nonetheless, the description of the distribution P_G of the generative models, which involves the description of the generated scattered data points [8] based on the distribution P_X of the original manifold $\mathcal{X}$, is very difficult using traditional distance measures, such as the Fréchet Inception Distance [7]. We highlight the distribution and the manifold notations in figure 1 for GAN and in figure 2 for AE. Effectively, traditional distance measures are not able to acknowledge the shapes of the data manifolds and the scale at which the manifold should be analyzed. However, persistent homology [9], [10] is specifically designed to highlight the topological features of the data [11]. Therefore, building upon persistent homology, Wasserstein distance [12] and generative models [7], our main contribution is to propose qualitative and quantitative ways to evaluate the scattered generated distributions and the performance of the generative models.

In this work we describe the persistent homology features of the generated model G while minimizing the OT function $W_c(P_X, P_G)$ for a squared cost $c(x, y) = ||x - y||_2^2$ where P_X is the model distribution of the data contained in the manifold $\mathcal{X}$, and P_G the distribution of the generative model capable of generating adversarial samples. Our contributions are summarized below:

- A persistent homology procedure for generative models, including GP-WGAN, WGAN, WAE and VAE, which we call PHom-GeM to highlight the topological properties of the generated distributions of the data for different spatial resolutions. The objective is a persistent homology de-

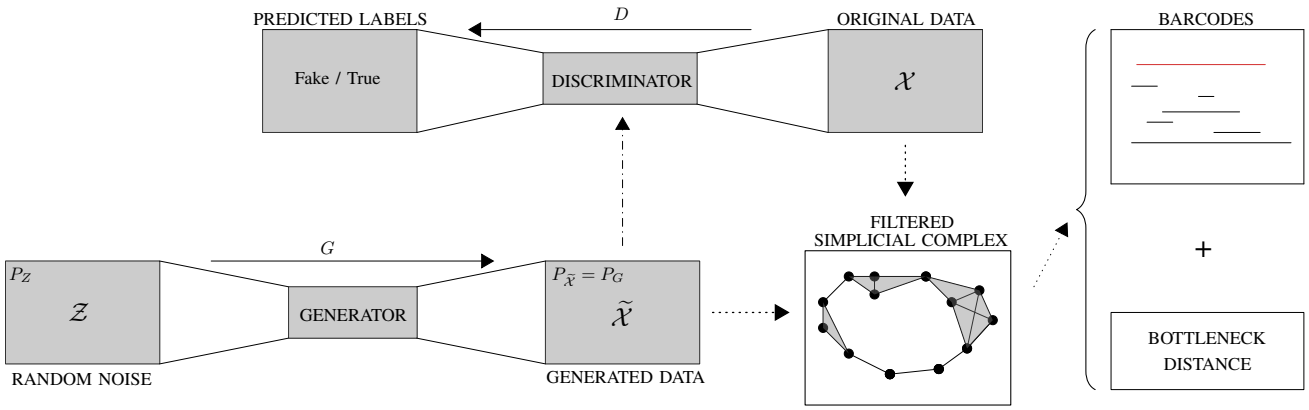

Fig. 1: In PHom-GeM applied to GAN, the generative model G generates fake samples $\widetilde{X} \in \widetilde{\mathcal{X}}$ based on the samples $Z \in \mathcal{Z}$ from the prior random distribution P_Z. Then, the discriminator model D tries to differentiate the fake samples $\widetilde{X}$ from the true samples $X \in \mathcal{X}$. The original manifold $\mathcal{X}$ and the generated manifold $\widetilde{\mathcal{X}}$ are transformed independently into metric space sets to obtain filtered simplicial complex. It leads to the description of topological features, summarized by the barcodes, to compare the respective topological representation of the true data distribution P_X and the generative model distribution P_G.

scription of the generated data distribution P_G following the generative model G.

- A distance measure for persistence diagrams, the bottleneck distance, applied to generative models to compare quantitatively the true and the target distributions on any data set. We measure the shortest distance for which there exists a perfect matching between the points of the two persistence diagrams. A persistence diagram is a stable summary representation of topological features of simplicial complex, a collection of vertices, associated to the data set.
- Finally, we propose the first application of algebraic topology and generative models on a public data set containing credit card transactions, particularly challenging for this type of models and traditional distance measures.

The paper is structured as follows. In section II, we review the optimized GP-WGAN and WAE formulations using OT derived by Gulrajani et al. in [6] and Tolstikhin et al. in [7], respectively. By using persistence homology, we are able to compare the topological properties of the original distribution P_X and the generated distribution P_G. We highlight experimental results in section III and we conclude in section IV by addressing promising directions for future work.

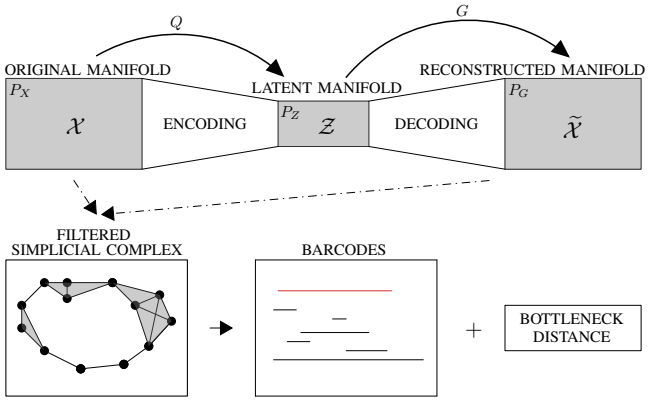

Fig. 2: In PHom-GeM applied to AE, the generative model G, the decoder, is used to generate fake samples $\widetilde{X} \in \widetilde{\mathcal{X}}$ based on the samples $Z \in \mathcal{Z}$ from a prior random distribution P_Z. Afterward, the original manifold $\mathcal{X}$ and the generated manifold $\widetilde{\mathcal{X}}$ are both transformed independently into metric space sets to obtain filtered simplicial complex. As for PHom-GeM applied to GAN, it leads to the description of topological features, summarized by the barcodes, to compare the respective topological representation of the true data distribution P_X and the generative model distribution P_G.

II. Proposed Method

Our method computes the persistence homology of both the true manifold $X \in \mathcal{X}$ and the generated manifold $\widetilde{X} \in \widetilde{\mathcal{X}}$ following the generative model G based on the minimization of the optimal transport cost $W_c(P_X, P_G)$. In the resulting topological problem, the points of the manifolds are transformed to a metric space set for which a Vietoris-Rips simplicial complex filtration is applied (see definition 2). PHom-GeM achieves simultaneously two main goals: it computes the birth-death of the pairing generators of the iterated inclusions while measuring the bottleneck distance between persistence diagrams of the manifolds of the generative models.

A. Optimal Transport and Dual Formulation

Following the description of the optimal transport problem [4] and relying on the Kantorovich-Rubinstein duality, the Wasserstein distance is computed as

$$W_c(P_X, P_G) = \sup_{f \in \mathcal{F}_L} \mathbb{E}_{X \sim P_X}[f(X)] - \mathbb{E}_{Y \sim P_G}[f(Y)] \quad (1)$$

where $(\mathcal{X}, d)$ is a metric space, $\mathcal{P}(X \sim P_X, Y \sim P_G)$ is a set of all joint distributions (X, Y) with marginals P_X and P_G respectively and $\mathcal{F}_L$ is the class of all bounded 1-Lipschitz functions on $(\mathcal{X}, d)$.

B. Gradient Penalty Wasserstein GAN (GP-WGAN)

As described in [6], the GP-WGAN objective loss function with gradient penalty is expressed such that

$$\begin{aligned} L = &\mathop{\mathbb{E}}_{\widetilde{X}\sim P_G}[f(\widetilde{X})] - \mathop{\mathbb{E}}_{X\sim P_X}[f(X)] \\ &+ \lambda \mathop{\mathbb{E}}_{\widehat{X}\sim P_{\widehat{X}}}[(||\nabla_{\widehat{X}} f(\widehat{X})||_2 - 1)^2] \end{aligned} \tag{2}$$

where f is the set of 1-Lipschitz functions on $(\mathcal{X}, d)$, P_X the original data distribution, P_G the generative model distribution implicitly defined by $\widetilde{X} = G(Z), Z \sim p(Z)$. The input Z to the generator is sampled from a noise distribution such as a uniform distribution. $P_{\widehat{X}}$ defines the uniform sampling along straight lines between pairs of points sampled from the data distribution P_X and the generative distribution P_G. A penalty on the gradient norm is enforced for random samples $\widehat{X} \sim P_{\widehat{X}}$. For further details, we refer to [6] and [5].

C. Wasserstein Auto-Encoders

As described in [7], the WAE objective function is expressed such that

$$\begin{aligned} D_{\text{WAE}}(P_X, P_G) := &\inf_{Q(Z|X)\in\mathcal{Q}} \mathbb{E}_{P_X}\mathbb{E}_{Q(Z|X)}[c(X, G(Z))] \\ &+ \lambda \mathcal{D}_Z(Q_Z, P_Z) \end{aligned} \tag{3}$$

where $c(X, G(Z)) : \mathcal{X} \times \mathcal{X} \to \mathcal{R}_+$ is any measurable cost function. In our experiments, we use a square cost function $c(x, y) = ||x - y||_2^2$ for data points $x, y \in \mathcal{X}$. $G(Z)$ denotes the sending of Z to X for a given map $G : \mathcal{Z} \to \mathcal{X}$. Q, and G, are any nonparametric set of probabilistic encoders, and decoders respectively.

We use the Maximum Mean Discrepancy (MMD) for the penalty $\mathcal{D}_Z(Q_Z, P_Z)$ for a positive-definite reproducing kernel $k : \mathcal{Z} \times \mathcal{Z} \to \mathcal{R}$

$$\begin{aligned} \mathcal{D}_Z(P_Z, Q_Z) :=&\text{MMD}_k(P_Z, Q_Z) \\ =&|| \int_{\mathcal{Z}} k(z, .)dP_Z(z) - \int_{\mathcal{Z}} k(z, .)dQ_Z(z)||_{\mathcal{H}_k} \end{aligned} \tag{4}$$

where $\mathcal{H}_k$ is the reproducing kernel Hilbert space of real-valued functions mapping on $\mathcal{Z}$. For details on the MMD implementation, we refer to [7].

D. Persistence Diagram and Vietoris-Rips Complex

Definition 1 Let $V = \{1, \cdots, |V|\}$ be a set of vertices. A simplex σ is a subset of vertices $\sigma \subseteq V$. A simplicial complex K on V is a collection of simplices $\{\sigma\}$, $\sigma \subseteq V$, such that $\tau \subseteq \sigma \in K \Rightarrow \tau \in K$. The dimension $n = |\sigma| - 1$ of σ is its number of elements minus 1. Simplicial complexes examples are represented in figure 3.

Definition 2 Let (X, d) be a metric space. The Vietoris-Rips complex $\text{VR}(X, \epsilon)$ at scale ϵ associated to X is the abstract simplicial complex whose vertex set is X, and where $\{x_0, x_1, ..., x_k\}$ is a k-simplex if and only if $d(x_i, x_j) \leq \epsilon$ for all $0 \leq i, j \leq k$.

We obtain an increasing sequence of Vietoris-Rips complex by considering the $\text{VR}(\mathcal{C}, \epsilon)$ for an increasing sequence $(\epsilon_i)_{1\leq i\leq N}$ of value of the scale parameter ϵ

$$\mathcal{K}_1 \overset{\iota}{\hookrightarrow} \mathcal{K}_2 \overset{\iota}{\hookrightarrow} \mathcal{K}_3 \overset{\iota}{\hookrightarrow} ... \overset{\iota}{\hookrightarrow} \mathcal{K}_{N-1} \overset{\iota}{\hookrightarrow} \mathcal{K}_N. \tag{5}$$

Applying the *k-th* singular homology functor $H_k(-, F)$ with coefficient in the field F [13] to (5), we obtain a sequence of F-vector spaces, called the *k-th* persistence module of $(\mathcal{K}_i)_{1\leq i\leq N}$

$$\begin{aligned} H_k(\mathcal{K}_1, F) \xrightarrow{t_1} H_k(\mathcal{K}_2, F) \xrightarrow{t_2} \cdots \xrightarrow{t_{N-2}} \\ H_k(\mathcal{K}_{N-1}, F) \xrightarrow{t_{N-1}} H_k(\mathcal{K}_N, F). \end{aligned} \tag{6}$$

Definition 3 $\forall\ i < j$, the *(i,j)*-persistent k-homology group with coefficient in F of $\mathcal{K} = (\mathcal{K}_i)_{1\leq i\leq N}$ denoted $HP_k^{i\to j}(\mathcal{K}, F)$ is defined to be the image of the homomorphism $t_{j-1} \circ \ldots \circ t_i : H_k(\mathcal{K}_i, F) \to H_k(\mathcal{K}_j, F)$.

Using the interval decomposition theorem [14], we extract a finite family of intervals of $\mathbb{R}_+$ called persistence diagram. Each interval can be considered as a point in the set $D = \{(x, y) \in \mathbb{R}_+^2 | x \leq y\}$. Hence, we obtain a finite subset of the set D. This space of finite subsets is endowed with a matching distance called the bottleneck distance and defined as follow

$$d_b(A, B) = \inf_{\phi: A'\to B'} \sup_{x\in A'} \|x - \phi(x)\|$$

where $A' = A \cup \Delta$, $B' = B \cup \Delta$, $\Delta = \{(x, y) \in \mathbb{R}_+^2 | x = y\}$ and the inf is over all the bijections from A' to B'.

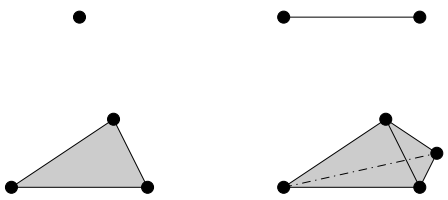

Fig. 3: A simplical complex is a collection of numerous "simplex" or simplices, where a 0-simplex is a point, a 1-simplex is a line segment, a 2-simplex is a triangle and a 3-simplex is a tetrahedron.

E. Application: PHom-GeM, Persistent Homology for Generative Models

Bridging the gap between persistent homology and generative models, PHom-GeM uses a two-steps procedure. First, the minimization problem is solved for the generator G and the discriminator D when considering GP-WGAN and WGAN. The gradient penalty λ in equation (2) is fixed equal to 10 for GP-WGAN and to 0 for WGAN. For auto-encoders, the minimization problem is solved for the encoder Q and the decoder G. We use RMSProp optimizer [15] for the optimization procedure. Then, the samples of the original and generated distributions, P_X and P_G, are mapped to persistence homology for the description of their respective manifolds. The points contained in the manifold $\mathcal{X}$ inherited from P_X and the points contained in the manifold $\tilde{\mathcal{X}}$ generated with P_G are randomly selected into respective batches. Two samples, Y_1 from $\mathcal{X}$ following P_X and Y_2 from $\tilde{\mathcal{X}}$ following P_G, are selected to differentiate the topological features of the original manifold $\mathcal{X}$ and the generated manifold $\tilde{\mathcal{X}}$. The samples Y_1 and Y_2 are contained in the spaces $\mathcal{Y}_1$ and $\mathcal{Y}_2$, respectively. Then, the spaces $\mathcal{Y}_1$ and $\mathcal{Y}_2$ are transformed into metric space sets $\widehat{\mathcal{Y}}_1$ and $\widehat{\mathcal{Y}}_2$ for computational purposes. Then, we filter the metric space sets $\widehat{\mathcal{Y}}_1$ and $\widehat{\mathcal{Y}}_2$ using the Vietoris-Rips simplicial complex filtration. Given a line segment of length ϵ, vertices between data points are created for data points respectively separated from a smaller distance than ϵ. It leads to the construction of a collection of simplices resulting in Vietoris-Rips simplicial complex $\text{VR}(\mathcal{C}, \epsilon)$ filtration. In our case, we decide to use the Vietoris-Rips simplicial complex as it offers the best compromise between the filtration accuracy and the memory requirement [11]. Subsequently, the persistence diagrams, dgm_{Y_1} and dgm_{Y_2}, are constructed. We recall a persistence diagram is a stable summary representation of topological features of simplicial complex. The persistence diagrams allow the computation of the bottleneck distance $d_b(\text{dgm}_{Y_1}, \text{dgm}_{Y_2})$. Finally, the barcodes represent in a simple way the birth-death of the pairing generators of the iterated inclusions detected by the persistence diagrams.

Algorithm 1: Persistent Homology for Generative Models

Data: training and validation sets, hyperparameter λ
Result: persistent homology description of generative manifolds

begin
- */*Step 1: Generative Models Resolution*/*
- Select samples $\{x_1, ..., x_n\}$ from training set
- Select samples $\{z_1, ..., z_n\}$ from validation set
- With RMSProp gradient descent update ($\text{lr} = 0.001, \rho = 0.9, \epsilon = 10^{-6}$), optimize until convergence Q and G
  - **case** *GP-WGAN and WGAN:* using equation 2
  - **case** *WAE:* using equation 3
  - **case** *VAE:* using equation described in [3]
- */*Step 2: Persistence Diagram and Bottleneck Distance on manifolds of generative models*/*
- Random selection of samples $Y_1 \in \mathcal{Y}_1, Y_2 \in \mathcal{Y}_2$ from P_X and P_G
- Transform $\mathcal{Y}_1$ and $\mathcal{Y}_2$ spaces into a metric space set
- Filter the metric space set with a Vietoris-Rips simplicial complex $\text{VR}(\mathcal{C}, \epsilon)$
- Compute the persistence diagrams dgm_{Y_1} and dgm_{Y_2}
- Evaluate the bottleneck distance $d_b(\text{dgm}_{Y_1}, \text{dgm}_{Y_2})$
- Build the barcodes with respect to Y_1 and Y_2

return

III. Experiments

We empirically evaluate the proposed methodology PHom-GeM. We assess on a highly challenging data set for generative models whether PHom-GeM can simultaneously achieve (i) precise persistent homology mapping of the generated data points and (ii) accurate persistent homology distance measurement with the bottleneck distance.

Data Availability and Data Description We train PHom-GeM on one real-world open data set: the credit card transactions data set from the Kaggle database[1] containing 284 807 transactions including 492 frauds. This data set is particularly interesting because it reflects the scattered points distribution of the reconstructed manifold that are found during generative models' training, impacting afterward the generated adversarial samples. Furthermore, this data set is challenging because of the strong imbalance between normal and fraudulent transactions while being of high interest for the banking industry. To preserve transactions confidentiality, each transaction is composed of 28 components obtained with PCA without any description and two additional features *Time* and *Amount* that remained unchanged. Each transaction is labeled as fraudulent or normal in a feature called *Class* which takes a value of 1 or 0, respectively.

Experimental Setup and Code Availability In our experiments, we use the Euclidean latent space $\mathcal{Z} = \mathcal{R}^2$ and the square cost function c previously defined as $c(x, y) = ||x - y||_2^2$ for the data points $x \in \mathcal{X}, \widetilde{x} \in \widetilde{\mathcal{X}}$. The dimensions of the true data set is $\mathcal{R}^{29}$. We kept the 28 components obtained with PCA and the amount resulting in a space of dimension 29. For the error minimization process, we used RMSProp gradient descent [15] with the parameters $\text{lr} = 0.001, \rho = 0.9, \epsilon = 10^{-6}$ and a batch size of 64. Different values of λ for the gradient penalty have been tested. We empirically obtained the lowest error reconstruction with $\lambda = 10$ for both GP-WGAN and WAE. The coefficients of persistence homology are evaluated within the field $\mathbb{Z}/2\mathbb{Z}$. We only consider homology groups H_0 and H_1 who represent the connected components and the loops, respectively. Higher dimensional homology groups did not noticeably improve the results quality while leading to longer computational time. The simulations were performed on a

[1]The data set is available at https://www.kaggle.com/mlg-ulb/creditcardfraud.

computer with 16GB of RAM, Intel i7 CPU and a Tesla K80 GPU accelerator. To ensure the reproducibility of the experiments, the code is available at the following address[2].

Results and Discussions about PHom-GeM We test PHom-GeM, Persistent Homology for Generative Models, on four different generative models: GP-WGAN, WGAN, WAE and VAE. We compare the performance of PHom-GeM on two specificities: first, qualitative visualization of the persistence diagrams and barcodes and, secondly, quantitative estimation of the persistent homology closeness using the bottleneck distance between the generated manifolds $\widetilde{\mathcal{X}}$ of the generative models and the original manifold $\mathcal{X}$.

On the top of figure 4, the rotated persistence and the barcode diagrams of the original sample $\mathcal{X}$ are highlighted. In the persistence diagram, black points represent the 0-dimensional homology groups H_0, the connected components of the complex. The red triangles represent the 1-dimensional homology group H_1, the 1-dimensional features known as cycles or loops. The barcode diagram is a simple way of representing the information contained in the persistence diagram. For the sake of simplicity, we represent only the barcode diagram of the generative models to compare qualitatively the generated distribution P_G of each model with respect to the distribution P_X of the original sample. The generated distribution P_G of GP-WGAN is the closest to the distribution P_X followed by WGAN, WAE and VAE. Effectively, the spectrum of the barcodes of GP-WGAN is very similar to the original sample's spectrum as well as denser on the right. On the opposite, the WAE and VAE's distributions P_G are not able to reproduce all of the features contained in the original distribution, therefore explaining the narrower barcode spectrum.

In order to quantitatively assess the quality of the generated distributions, we use the bottleneck distance between the persistent diagram of $\mathcal{X}$ and the persistent diagram of $G(Z)$ of the generated data points. In table I, we highlight the mean value of the bottleneck distance for a 95% confidence interval. We also underline the lower and the upper bounds of the 95% confidence interval for each generative model. Confirming the visual observations, we notice the smallest bottleneck distance, and therefore, the best result, is obtained with GP-WGAN, followed by WGAN, WAE and VAE. It means GP-WGAN is capable to generate data distribution sharing the most topological features with the original data distribution, including the nearness measurements and the overall shape. It confirms topologically on a real-world data set the claims addressed in [6] of superior performance of GP-WGAN against WGAN. Furthermore, the performance of the AE cannot match the generative performance achieved by the GANs. However, the WAE, that relies on optimal transport theory, achieves better generative distribution in comparison to the popular VAE.

[2]The code is available at https://github.com/dagrate/phomgem

TABLE I: Bottleneck distance (smaller is better) with 95% of confidence interval between the samples X of the original manifold $\mathcal{X}$ and the generated samples $\widetilde{X}$ of the manifold $\widetilde{\mathcal{X}}$. Because of the Wasserstein distance and gradient penalty, GP-WGAN achieves better performance.

Gen. Model	Mean Value	Lower Bound	Upper Bound
GP-WGAN	**0.0711**	**0.0683**	**0.0738**
WGAN	0.0744	0.0716	0.0772
WAE	0.0821	0.0791	0.0852
VAE	0.0857	0.0833	0.0881

IV. Conclusion

Building upon optimal transport and unsupervised learning, we introduced PHom-GeM, Persistent Homology for Generative Models, a new characterization of the generative manifolds that uses topology and persistence homology to highlight manifold features and scattered generated distributions. We discuss the relations of GP-WGAN, WGAN, WAE and VAE in the context of unsupervised learning. Furthermore, relying on persistence homology, the bottleneck distance has been introduced to estimate quantitatively the topological features similarities between the original distribution and the generated distributions of the generative models, a specificity that current traditional distance measures fail to acknowledge. We conducted experiments showing the performance of PHom-GeM on the four generative models GP-WGAN, WGAN, WAE and VAE. We used a challenging imbalanced real-world open data set containing credit card transactions, capable of illustrating the scattered generated data distributions of the generative models and particularly suitable for the banking industry. We showed the superior topological performance of GP-WGAN in comparison to the other generative models as well as the superior performance of WAE over VAE. Future work will include further exploration of the topological features such as the influence of the simplicial complex and the possibility to integrate a topological optimization function as a regularization term.

References

[1] Ian Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, and Yoshua Bengio. Generative adversarial nets. In *Advances in neural information processing systems*, pages 2672–2680, 2014.

[2] Ian Goodfellow, Yoshua Bengio, Aaron Courville, and Yoshua Bengio. *Deep learning*, volume 1. MIT press Cambridge, 2016.

[3] Diederik P Kingma and Max Welling. Auto-encoding variational bayes. *arXiv preprint arXiv:1312.6114*, 2013.

[4] Cédric Villani. *Topics in optimal transportation*. Number 58. American Mathematical Soc., 2003.

[5] Martin Arjovsky, Soumith Chintala, and Léon Bottou. Wasserstein generative adversarial networks. In *International Conference on Machine Learning*, pages 214–223, 2017.

[6] Ishaan Gulrajani, Faruk Ahmed, Martin Arjovsky, Vincent Dumoulin, and Aaron C Courville. Improved training of wasserstein gans. In *Advances in Neural Information Processing Systems*, pages 5767–5777, 2017.

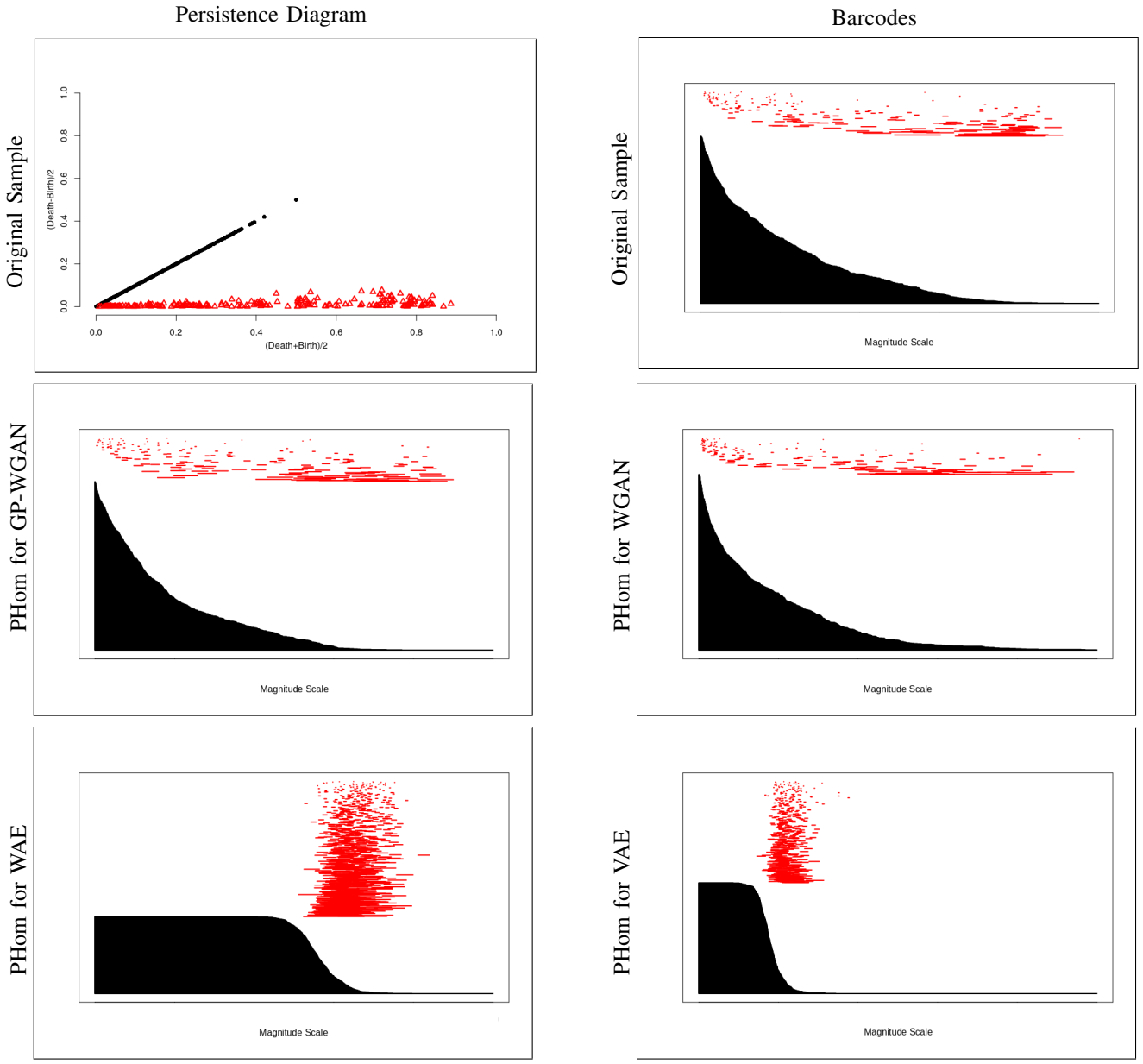

Fig. 4: On top, the rotated persistence and the barcode diagrams of the original sample are represented. They both illustrate the birth-death of the pairing generators of the iterated inclusions. In the persistence diagram, the black points represent the connected components of the complex and the red triangles the cycles. Below, GP-WGAN, WGAN, WAE and VAE's barcode diagrams allows to assess qualitatively, with the original sample barcode diagram, the persistent homology similarities between the generated and the original distribution, P_G and P_X respectively.

[7] Ilya Tolstikhin, Olivier Bousquet, Sylvain Gelly, and Bernhard Schoelkopf. Wasserstein auto-encoders. *arXiv preprint arXiv:1711.01558*, 2017.

[8] Yoshua Bengio, Li Yao, Guillaume Alain, and Pascal Vincent. Generalized denoising auto-encoders as generative models. In *Advances in Neural Information Processing Systems*, pages 899–907, 2013.

[9] Herbert Edelsbrunner, David Letscher, and Afra Zomorodian. Topological Persistence and Simplification. *Discrete & Computational Geometry*, 28(4):511–533, 2002.

[10] Afra Zomorodian and Gunnar Carlsson. Computing persistent homology. *Discrete & Computational Geometry*, 33(2):249–274, 2005.

[11] Frédéric Chazal and Bertrand Michel. An introduction to Topological Data Analysis: fundamental and practical aspects for data scientists. *arXiv preprint arXiv:1710.04019*, 2017.

[12] Olivier Bousquet, Sylvain Gelly, Ilya Tolstikhin, Carl-Johann Simon-Gabriel, and Bernhard Schoelkopf. From optimal transport to generative modeling: the vegan cookbook. *arXiv preprint arXiv:1705.07642*, 2017.

[13] A. Hatcher. *Algebraic Topology*. Cambridge University Press, 2002.

[14] Steve Y. Oudot. *Persistence Theory: From Quiver Representations to Data Analysis*. Number 209 in Mathematical Surveys and Monographs. American Mathematical Society, 2015.

[15] Geoffrey Hinton, Nitish Srivastava, and Kevin Swersky. Rmsprop: Divide the gradient by a running average of its recent magnitude. *Neural networks for machine learning, Coursera lecture 6e*, 2012.